

# KMeans\_Clustering

April 26, 2018

## 1 Code for number of homicides

```
In [2]: %matplotlib inline
        from copy import deepcopy
        import numpy as np
        import pandas as pd
        from matplotlib import pyplot as plt
        plt.rcParams['figure.figsize'] = (16, 9)
        plt.style.use('ggplot')

        #=====Plot for number of homicides=====
        numberHomicides = []
        for i in range(2001,2018):
            print('Reading file '+str(i)+'...')
            fileName = 'day_type_description_crap_'+str(i)+'.csv'
            plotDf = pd.read_csv(fileName)
            n = plotDf['PrimaryType HOMICIDE'].sum()
            numberHomicides.append({'Year': i, 'numberHomicides': n})

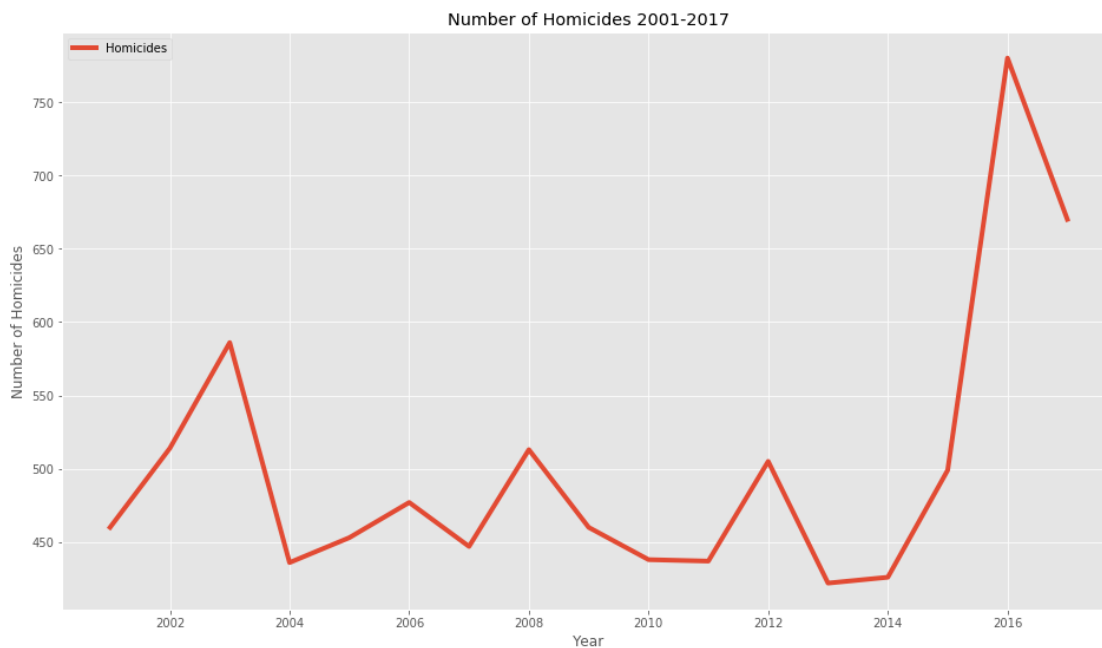
        homicidesDf = pd.DataFrame(numberHomicides)
        plt.plot(homicidesDf['Year'],homicidesDf['numberHomicides'], linewidth=4, label='Homicides')
        plt.xlabel('Year')
        plt.ylabel('Number of Homicides')
        plt.title('Number of Homicides 2001-2017')
        plt.legend()
        plt.savefig('numberHomicides.png')

        Reading file 2001...
        Reading file 2002...
        Reading file 2003...
        Reading file 2004...
        Reading file 2005...
        Reading file 2006...
        Reading file 2007...
        Reading file 2008...
        Reading file 2009...
        Reading file 2010...
```

```

Reading file 2011...
Reading file 2012...
Reading file 2013...
Reading file 2014...
Reading file 2015...
Reading file 2016...
Reading file 2017...

```



## 2 Attempt to parse data and store it to use in kmeans clustering

```

In [ ]: df = pd.read_csv('day_type_description_crap_2001.csv')
del df['Unnamed: 0']      #Deleting first column
del df['YearMonthDay']    #Deleting second column
del df['ID']              #Deleting third column
df['CommunityArea'] = df['CommunityArea'].astype(int)

df2 = pd.read_csv('day_type_description_crap_2002.csv')
del df2['Unnamed: 0']     #Deleting first column
del df2['YearMonthDay']   #Deleting second column
del df2['ID']             #Deleting third column
df2['CommunityArea'] = df2['CommunityArea'].astype(int)

df3 = pd.read_csv('day_type_description_crap_2003.csv')
del df3['Unnamed: 0']     #Deleting first column
del df3['YearMonthDay']   #Deleting second column

```

```

del df3['ID'] #Deleting third column
df3['CommunityArea'] = df3['CommunityArea'].astype(int)

df4 = pd.read_csv('day_type_description_crap_2004.csv')
del df4['Unnamed: 0'] #Deleting first column
del df4['YearMonthDay'] #Deleting second column
del df4['ID'] #Deleting third column
df4['CommunityArea'] = df4['CommunityArea'].astype(int)

df5 = pd.read_csv('day_type_description_crap_2005.csv')
del df5['Unnamed: 0'] #Deleting first column
del df5['YearMonthDay'] #Deleting second column
del df5['ID'] #Deleting third column
df5['CommunityArea'] = df5['CommunityArea'].astype(int)

df6 = pd.read_csv('day_type_description_crap_2006.csv')
del df6['Unnamed: 0'] #Deleting first column
del df6['YearMonthDay'] #Deleting second column
del df6['ID'] #Deleting third column
df6['CommunityArea'] = df6['CommunityArea'].astype(int)

df7 = pd.read_csv('day_type_description_crap_2007.csv')
del df7['Unnamed: 0'] #Deleting first column
del df7['YearMonthDay'] #Deleting second column
del df7['ID'] #Deleting third column
df7['CommunityArea'] = df7['CommunityArea'].astype(int)

df8 = pd.read_csv('day_type_description_crap_2008.csv')
del df8['Unnamed: 0'] #Deleting first column
del df8['YearMonthDay'] #Deleting second column
del df8['ID'] #Deleting third column
df8['CommunityArea'] = df8['CommunityArea'].astype(int)

df9 = pd.read_csv('day_type_description_crap_2009.csv')
del df9['Unnamed: 0'] #Deleting first column
del df9['YearMonthDay'] #Deleting second column
del df9['ID'] #Deleting third column
df9['CommunityArea'] = df9['CommunityArea'].astype(int)

df10 = pd.read_csv('day_type_description_crap_2010.csv')
del df10['Unnamed: 0'] #Deleting first column
del df10['YearMonthDay'] #Deleting second column
del df10['ID'] #Deleting third column
df10['CommunityArea'] = df10['CommunityArea'].astype(int)

df11 = pd.read_csv('day_type_description_crap_2011.csv')
del df11['Unnamed: 0'] #Deleting first column
del df11['YearMonthDay'] #Deleting second column

```

```

del df11['ID']                                #Deleting third column
df11['CommunityArea'] = df11['CommunityArea'].astype(int)

df12 = pd.read_csv('day_type_description_crap_2012.csv')
del df12['Unnamed: 0']                        #Deleting first column
del df12['YearMonthDay']                     #Deleting second column
del df12['ID']                               #Deleting third column
df12['CommunityArea'] = df12['CommunityArea'].astype(int)

df13 = pd.read_csv('day_type_description_crap_2013.csv')
del df13['Unnamed: 0']                        #Deleting first column
del df13['YearMonthDay']                     #Deleting second column
del df13['ID']                               #Deleting third column
df13['CommunityArea'] = df13['CommunityArea'].astype(int)

df14 = pd.read_csv('day_type_description_crap_2014.csv')
del df14['Unnamed: 0']                        #Deleting first column
del df14['YearMonthDay']                     #Deleting second column
del df14['ID']                               #Deleting third column
df14['CommunityArea'] = df14['CommunityArea'].astype(int)

df15 = pd.read_csv('day_type_description_crap_2015.csv')
del df15['Unnamed: 0']                        #Deleting first column
del df15['YearMonthDay']                     #Deleting second column
del df15['ID']                               #Deleting third column
df15['CommunityArea'] = df15['CommunityArea'].astype(int)

df16 = pd.read_csv('day_type_description_crap_2016.csv')
del df16['Unnamed: 0']                        #Deleting first column
del df16['YearMonthDay']                     #Deleting second column
del df16['ID']                               #Deleting third column
df16['CommunityArea'] = df16['CommunityArea'].astype(int)

df17 = pd.read_csv('day_type_description_crap_2017.csv')
del df17['Unnamed: 0']                        #Deleting first column
del df17['YearMonthDay']                     #Deleting second column
del df17['ID']                               #Deleting third column
df17['CommunityArea'] = df17['CommunityArea'].astype(int)

crimeAttributes = list(df)
del crimeAttributes[0]                        #Deleting "CommunityArea" attribute
del crimeAttributes[0]                        #Deleting "IndexCrime" attribute

comunCrime = []
for i in range(1,78):
    tempDF = df.loc[df['CommunityArea'] == i]
    tempDF2 = df2.loc[df2['CommunityArea'] == i]

```

```

tempDF3 = df3.loc[df3['CommunityArea'] == i]
tempDF4 = df4.loc[df4['CommunityArea'] == i]
tempDF5 = df5.loc[df5['CommunityArea'] == i]
tempDF6 = df6.loc[df6['CommunityArea'] == i]
tempDF7 = df7.loc[df7['CommunityArea'] == i]
tempDF8 = df8.loc[df8['CommunityArea'] == i]
tempDF9 = df9.loc[df9['CommunityArea'] == i]
tempDF10 = df10.loc[df10['CommunityArea'] == i]
tempDF11 = df11.loc[df11['CommunityArea'] == i]
tempDF12 = df12.loc[df12['CommunityArea'] == i]
tempDF13 = df13.loc[df13['CommunityArea'] == i]
tempDF14 = df14.loc[df14['CommunityArea'] == i]
tempDF15 = df15.loc[df15['CommunityArea'] == i]
tempDF16 = df16.loc[df16['CommunityArea'] == i]
tempDF17 = df17.loc[df17['CommunityArea'] == i]

totalCrimes = 0
totalCrimes2 = 0
totalCrimes3 = 0
totalCrimes4 = 0
totalCrimes5 = 0
totalCrimes6 = 0
totalCrimes7 = 0
totalCrimes8 = 0
totalCrimes9 = 0
totalCrimes10 = 0
totalCrimes11 = 0
totalCrimes12 = 0
totalCrimes13 = 0
totalCrimes14 = 0
totalCrimes15 = 0
totalCrimes16 = 0
totalCrimes17 = 0

for j in range(len(crimeAttributes)):
    totalCrimes += tempDF[crimeAttributes[j]].sum()
    totalCrimes2 += tempDF2[crimeAttributes[j]].sum()
    totalCrimes3 += tempDF3[crimeAttributes[j]].sum()
    totalCrimes4 += tempDF4[crimeAttributes[j]].sum()
    totalCrimes5 += tempDF5[crimeAttributes[j]].sum()
    totalCrimes6 += tempDF6[crimeAttributes[j]].sum()
    totalCrimes7 += tempDF7[crimeAttributes[j]].sum()
    totalCrimes8 += tempDF8[crimeAttributes[j]].sum()
    totalCrimes9 += tempDF9[crimeAttributes[j]].sum()
    totalCrimes10 += tempDF10[crimeAttributes[j]].sum()
    totalCrimes11 += tempDF11[crimeAttributes[j]].sum()
    totalCrimes12 += tempDF12[crimeAttributes[j]].sum()
    totalCrimes13 += tempDF13[crimeAttributes[j]].sum()

```

```

totalCrimes14 += tempDF14[crimeAttributes[j]].sum()
totalCrimes15 += tempDF15[crimeAttributes[j]].sum()
totalCrimes16 += tempDF16[crimeAttributes[j]].sum()
totalCrimes17 += tempDF17[crimeAttributes[j]].sum()

comunCrime.append({'communityArea':i, 'totalCrimes-1':totalCrimes, 'totalCrimes-2':
                    'totalCrimes-3':totalCrimes3, 'totalCrimes-4':totalCrimes4, 'totalCrimes-5':totalCrimes5,
                    'totalCrimes-6':totalCrimes6, 'totalCrimes-7':totalCrimes7, 'totalCrimes-8':totalCrimes8,
                    'totalCrimes-9':totalCrimes9, 'totalCrimes-10':totalCrimes10, 'totalCrimes-11':totalCrimes11,
                    'totalCrimes-12':totalCrimes12, 'totalCrimes-13':totalCrimes13, 'totalCrimes-14':totalCrimes14,
                    'totalCrimes-15':totalCrimes15, 'totalCrimes-16':totalCrimes16, 'totalCrimes-17':totalCrimes17})

display(updatedDf)
updatedDf = pd.DataFrame(comunCrime)
updatedDf.to_csv('Updated.csv')

```

### 3 Final code to parse data for kmeans clustering

```

In [24]: updatedDf = pd.DataFrame()
comunList = list()
for i in range(1,78):
    comunList.append(i)

updatedDf['CommunityArea'] = comunList

counter = 1;
for i in range(2001,2018):
    comunCrimes = list()
    print('Reading file '+str(i)+'...')
    fileName = 'day_type_description_crap_'+str(i)+'.csv'
    df = pd.read_csv(fileName)
    del df['Unnamed: 0']          #Deleting first column
    del df['YearMonthDay']       #Deleting second column
    del df['ID']                 #Deleting third column
    df['CommunityArea'] = df['CommunityArea'].astype(int)

    crimeAttributes = list(df)
    del crimeAttributes[0]       #Deleting "CommunityArea" attribute
    del crimeAttributes[0]       #Deleting "IndexCrime" attribute

    print('\t Processing data from file: '+fileName+'\n')
    for k in range(1,78):
        tempDF = df.loc[df['CommunityArea'] == k]
        totalCrimes = 0
        for j in range(len(crimeAttributes)):
            totalCrimes += tempDF[crimeAttributes[j]].sum()
        comunCrimes.append(totalCrimes)

```

```

        updatedDf['totalCrimes'+str(counter)] = comunCrimes
        counter += 1

allCrimes = list()
for i in range(0,77):
    allCrimes.append((updatedDf.iloc[i].sum())-(i+1))
updatedDf['sumAllCrimes'] = allCrimes

print('\t \n ===== FINISHED ===== \n')
display(updatedDf)
updatedDf.to_csv('kmeans_data.csv')

summedData = pd.DataFrame()

Reading file 2001...
    Processing data from file: day_type_description_crap_2001.csv

Reading file 2002...
    Processing data from file: day_type_description_crap_2002.csv

Reading file 2003...
    Processing data from file: day_type_description_crap_2003.csv

Reading file 2004...
    Processing data from file: day_type_description_crap_2004.csv

Reading file 2005...
    Processing data from file: day_type_description_crap_2005.csv

Reading file 2006...
    Processing data from file: day_type_description_crap_2006.csv

Reading file 2007...
    Processing data from file: day_type_description_crap_2007.csv

Reading file 2008...
    Processing data from file: day_type_description_crap_2008.csv

Reading file 2009...
    Processing data from file: day_type_description_crap_2009.csv

Reading file 2010...
    Processing data from file: day_type_description_crap_2010.csv

Reading file 2011...
    Processing data from file: day_type_description_crap_2011.csv

Reading file 2012...

```

```

Processing data from file: day_type_description_crap_2012.csv

Reading file 2013...
Processing data from file: day_type_description_crap_2013.csv

Reading file 2014...
Processing data from file: day_type_description_crap_2014.csv

Reading file 2015...
Processing data from file: day_type_description_crap_2015.csv

Reading file 2016...
Processing data from file: day_type_description_crap_2016.csv

Reading file 2017...
Processing data from file: day_type_description_crap_2017.csv

```

===== FINISHED =====

	CommunityArea	totalCrimes1	totalCrimes2	totalCrimes3	totalCrimes4	\
0	1	56	6106	7684	7609	
1	2	41	4048	5235	5757	
2	3	72	6039	7522	7445	
3	4	33	2616	3271	3358	
4	5	25	2411	3277	3134	
5	6	52	6587	8807	8653	
6	7	57	6035	7948	7308	
7	8	95	12546	15470	15759	
8	9	4	323	427	400	
9	10	11	1466	2086	1894	
10	11	18	1440	1868	1707	
11	12	6	618	841	775	
12	13	18	1156	1473	1509	
13	14	39	3201	4434	3801	
14	15	51	4651	6508	5929	
15	16	51	4192	5495	5160	
16	17	30	2298	3078	2801	
17	18	11	924	1287	1036	
18	19	96	6399	8599	8144	
19	20	33	2105	3092	2842	
20	21	59	3607	4805	4827	
21	22	130	8918	11366	10652	
22	23	228	11236	15493	14541	
23	24	122	12341	15388	14156	
24	25	282	21112	30834	29487	



25	26	108	5595	7420	7880
26	27	112	6358	9179	9191
27	28	101	10853	15807	14768
28	29	147	8719	12348	13369
29	30	78	5917	7546	7892
..	...	...	...	...	...
47	48	24	1758	2463	2552
48	49	84	8800	11688	11808
49	50	18	1358	2010	1920
50	51	25	2129	2875	2636
51	52	13	1719	2557	2408
52	53	55	5540	7291	7888
53	54	14	1995	2346	2258
54	55	10	722	1066	1003
55	56	39	3620	4467	4521
56	57	14	1163	1632	1619
57	58	46	3379	4886	4723
58	59	17	1590	2084	2080
59	60	22	2422	3241	3127
60	61	73	7813	10822	10118
61	62	11	1359	1797	1714
62	63	43	3090	4272	3901
63	64	25	1452	1948	1912
64	65	29	2677	3534	3560
65	66	82	8135	11204	10976
66	67	83	10186	13919	14120
67	68	103	9544	12546	12959
68	69	82	7910	10905	11592
69	70	38	3003	3815	3962
70	71	96	8977	12300	13148
71	72	19	1447	1721	1743
72	73	45	4012	5084	5334
73	74	10	831	1129	1097
74	75	25	2768	3575	3635
75	76	33	1731	2218	2095
76	77	48	3771	4787	4897

	totalCrimes5	totalCrimes6	totalCrimes7	totalCrimes8	totalCrimes9 \
0	7530	7207	6807	6497	5959
1	5841	6052	5269	5470	5229
2	7543	6806	6395	6001	5188
3	3374	3285	2845	2784	2639
4	2990	2822	2634	2593	2432
5	7740	7842	8003	8215	7761
6	6618	6665	6915	6821	6118
7	14643	14222	13882	13513	12278
8	461	437	419	473	386
9	1748	1822	1780	2002	1714

10	1610	1804	1621	1777	1509
11	873	786	745	798	751
12	1600	1558	1502	1542	1311
13	3920	4093	3706	3821	3582
14	5855	5445	5345	5731	5294
15	5163	5067	4707	4909	4661
16	2653	2487	2693	2943	2696
17	1006	1004	1063	1018	1098
18	8367	7883	8034	8055	8316
19	2641	2740	2629	2486	2472
20	4356	4168	3885	3950	3874
21	9436	9600	9195	8928	8694
22	14800	13714	13056	13265	12772
23	12866	13013	12342	12729	11794
24	28376	28895	28535	27260	26036
25	8428	8414	7936	7427	6395
26	8317	8615	8412	7924	7169
27	13213	12147	11613	11590	10363
28	12737	12079	11347	11411	10426
29	7518	7438	7699	7341	6326
..	...	...	...	...	...
47	2339	2359	2349	2294	2111
48	11675	12273	12618	12284	11326
49	1820	1691	1952	1702	1487
50	2523	2571	2740	2694	2578
51	2291	2203	2099	2174	2091
52	7939	7660	8049	7139	6755
53	2008	1619	1529	1397	1402
54	928	965	859	971	819
55	4107	4214	3992	3658	3278
56	1614	1648	1731	1584	1651
57	4439	4647	4214	4175	3891
58	1908	2003	1788	1794	1759
59	3131	2975	2608	2659	2543
60	9648	10063	9725	8974	8199
61	1620	1744	1684	1727	1621
62	4327	4104	3917	3971	3947
63	1822	1824	1931	1907	1661
64	3233	3558	3209	3274	3097
65	11517	11737	11335	11266	10456
66	13674	13710	13561	13540	11528
67	12197	12311	12395	12437	10721
68	11207	10969	11283	10501	9695
69	4029	4305	4234	4315	3738
70	12818	12155	13327	12888	11769
71	1744	1529	1552	1656	1407
72	5336	5555	5412	5025	4719
73	1046	952	963	944	895

74	3563	3639	3655	3460	3372
75	2076	2519	2466	2356	2132
76	5015	4723	4246	4168	3613

	totalCrimes10	totalCrimes11	totalCrimes12	totalCrimes13	totalCrimes14	\
0	5577	5120	4925	4429	3801	
1	4882	4421	4029	3807	3409	
2	5074	4777	4976	4287	3810	
3	2560	2315	2338	2217	1897	
4	2062	1947	1920	1723	1540	
5	7459	7527	7696	6823	5695	
6	5795	5283	5010	4480	4086	
7	11594	11229	11285	10110	8987	
8	305	369	323	298	243	
9	1444	1447	1465	1230	1253	
10	1398	1300	1410	1336	1231	
11	695	605	600	443	461	
12	1188	1116	1082	917	872	
13	3547	3097	2945	2689	2400	
14	4639	4391	4160	3971	3386	
15	4529	4309	3937	3490	2847	
16	2350	2159	1930	1831	1568	
17	899	851	781	634	525	
18	7588	6872	6364	5708	4965	
19	2416	2062	2052	1965	1788	
20	3744	3372	3277	2673	2369	
21	8166	7331	6793	5852	5096	
22	11829	11151	11212	10255	9204	
23	11361	10139	9684	8207	7339	
24	24295	22715	21345	20229	18757	
25	6377	6281	6674	6676	6206	
26	6960	6412	6186	6272	5794	
27	9809	9668	9414	8507	7891	
28	10850	10008	10390	9790	8689	
29	6289	6019	5778	5107	4765	
..	...	...	...	...	...	
47	2098	2220	1923	1614	1588	
48	10378	9916	9026	8155	7374	
49	1374	1191	1262	1148	1219	
50	2622	2511	2293	2186	2253	
51	1854	1870	1746	1499	1386	
52	5977	5906	5748	5151	4368	
53	1396	1531	1586	1493	1382	
54	765	822	739	715	626	
55	2589	2554	2468	2107	1929	
56	1356	1286	1195	1200	1056	
57	3532	3712	3190	2895	2911	
58	1485	1429	1339	1278	1057	

59	2448	2329	2167	1891	1803
60	7739	7731	6814	5998	5258
61	1395	1487	1368	1292	1014
62	3945	3739	3239	2826	2624
63	1399	1339	1278	1115	999
64	2770	2681	2448	2251	2104
65	10516	9960	8590	7292	6482
66	11222	10823	9985	9013	7602
67	10164	9328	9092	8196	7105
68	9370	8924	8724	7726	6745
69	3473	3136	2823	2814	2569
70	10666	10262	9417	8576	7986
71	1286	1238	1205	1021	902
72	4548	4158	3963	3907	3194
73	916	727	707	626	671
74	3160	2908	2711	2492	2146
75	1894	1938	1832	2132	2012
76	3672	3338	3275	2715	2330

	totalCrimes15	totalCrimes16	totalCrimes17	sumAllCrimes
0	3584	3663	4122	90676
1	3105	3280	3565	73440
2	3656	3562	3536	86689
3	1770	1986	1980	41268
4	1393	1398	1413	35714
5	5587	6070	5698	116215
6	3758	4242	4447	91586
7	9062	10549	12233	197457
8	257	305	271	5701
9	1276	1320	1086	25044
10	1156	1071	1044	23300
11	449	572	544	10562
12	842	972	956	19614
13	2530	2412	2463	52680
14	3444	3238	3345	75383
15	3015	2986	2986	67504
16	1659	1753	1506	36435
17	587	670	608	14002
18	4829	4761	4569	109549
19	1731	1549	1499	36102
20	2261	2282	2329	55838
21	4801	5457	5036	125451
22	8178	8001	8027	186962
23	7075	8238	8269	175063
24	17339	16692	15305	377494
25	6019	5773	5346	108955
26	5442	5020	4728	112091
27	7933	9448	8997	172122

28	8383	8769	9001	168463
29	4448	4685	4497	99343
..	...	...	...	...
47	1482	1410	1352	31936
48	6627	6863	6905	157800
49	1047	1206	1089	23494
50	1920	1884	1746	38186
51	1316	1330	1206	29762
52	4002	3880	3975	97323
53	1186	1272	1344	25758
54	513	584	666	12773
55	1852	2104	1914	49413
56	996	879	840	21464
57	2756	2632	2201	58229
58	1083	1038	1048	24780
59	1613	1691	1494	38164
60	4985	4926	4619	123505
61	1035	1096	1013	22977
62	2364	2352	2291	54952
63	937	965	889	23403
64	2052	2137	1861	44475
65	6210	5817	5719	147294
66	7527	6990	6915	174398
67	7236	6276	5900	158510
68	6763	6591	6437	145424
69	2312	2718	2480	53764
70	7835	7585	7475	167280
71	988	983	882	21323
72	3158	3208	3065	69723
73	616	609	582	13321
74	2079	1987	2092	47267
75	1733	1769	1655	32591
76	2255	2382	2611	57846

[77 rows x 19 columns]

## 4 Failed attempt at Kmeans clustering

```
In [ ]: f1 = updatedDf['communityArea'].values
        f2 = updatedDf['totalCrimes'].values
        X=np.array(list(zip(f1,f2)))
        #Splitting into 4 smaller lists for plotting:
        comun1 = f1[:19]
        comun2 = f1[19:38]
        comun3 = f1[38:57]
        comun4 = f1[57:77]
```

```

totC1 = f2[:19]
totC2 = f2[19:38]
totC3 = f2[38:57]
totC4 = f2[57:77]

plt.scatter(comun1,totC1, color='red', label = 'communityAreas 1-19', s=100)
plt.scatter(comun2,totC2, color='green', label = 'communityAreas 20-38', s=100)
plt.scatter(comun3,totC3, color='yellow', label = 'communityAreas 39-57', s=100)
plt.scatter(comun4,totC4, color='blue', label = 'communityAreas 57-77', s=100)

legend = plt.legend(shadow=True)
frame = legend.get_frame()
frame.set_facecolor('1')
for label in legend.get_texts():
    label.set_fontsize('large')

```

In [ ]: # *Euclidean Distance Caculator*

```

from sklearn.cluster import KMeans

# Number of clusters
kmeans = KMeans(n_clusters=3)
# Fitting the input data
kmeans = kmeans.fit(X)
# Getting the cluster labels
labels = kmeans.predict(X)
# Centroid values
centroids = kmeans.cluster_centers_
print(centroids)

```

```

In [389]: c = centroids.ravel()
          C_x = c[[0,2,4]]
          C_y = c[[1,3,5]]
          plt.scatter(comun1,totC1, color='red', label = 'communityAreas 1-19', s=100)
          plt.scatter(comun2,totC2, color='green', label = 'communityAreas 20-38', s=100)
          plt.scatter(comun3,totC3, color='yellow', label = 'communityAreas 39-57', s=100)
          plt.scatter(comun4,totC4, color='blue', label = 'communityAreas 57-77', s=100)
          plt.scatter(C_x, C_y, marker='*', s=300, c='black')

```

Out[389]: <matplotlib.collections.PathCollection at 0x122a8cb00>

